between validity and utility seems to be merely the difference between more and less perfect forms of validity.

Jablensky's analysis makes the perfect the enemy of the good. He elevates the ideal goal of single-etiology validity to the only form of validity, when in fact validity is variegated, ranging from conceptual validity (successfully distinguishing normal vs. disordered conditions)[7], through many forms of partial validity (marking various homogeneities among the dysfunctions underlying a domain of disorders), to single-etiology construct validity. Then, once validity is placed out of reach (and pushed further out of reach by the embrace of a spurious zone-of-rarity validity criterion), he argues for an alternative focus on utility.

As my epigraph from a paper published in 1843[1] reflects, the idea that the difficulties in achieving validity in nosological classification should induce us to refocus on utility is not new. However, instead of "if it cannot be perfect, let it be useful", I would suggest the motto "if diagnostic criteria cannot be perfectly valid, let them be as valid as possible". That should be our goal, and in the long term it serves utility.

Recent deployments of utility as a nosological rationale on both sides of various disputes suggest as well a paraphrase of W. James's comment about the unconscious[8]: utility "is the sovereign means for believing what one likes in psychology, and of turning what might become a science into a tumbling ground for whimsies".

Jerome C. Wakefield
School of Social Work and Department of Psychiatry, New York University, New York, NY, USA

1. Johnson H. Prov Med J Retrosp Med Sci 1843; 6:472-3.
2. Jablensky A. World Psychiatry 2016;15:26-31.
3. Kendell R, Jablensky A. Am J Psychiatry 2003; 160:4-12.
4. Kendell RE, Brockington IF. Br J Psychiatry 1980;137:324-31.
5. Meehl PE. Psychodiagnosis. In: Selected papers. Minneapolis: University of Minnesota Press, 1959.
6. Klein DF. Mod Probl Pharmacopsychiatry 1987; 22:1-35.
7. Wakefield JC. Am Psychol 1992;47:73-88.
8. James W. Principles of psychology, Vol. 2. New York: Holt, 1890.

DOI:10.1002/wps.20286

# Utility without validity is useless

A. Jablensky's paper[1] raises important questions of many kinds. In this commentary, due to space constraints, I will not consider the evidence regarding clinical course as a diagnostic validator in schizophrenia or the overlap between psychosis and mood conditions, nor the nosological views of K. Jaspers, except to note that other interpretations exist that would not agree with Jablensky's perspective. The main focus here will be instead on whether the concept of utility can or should be the basis of psychiatric nosology.

The central assumption in Jablensky's paper is the statement, made in passing, that medical classifications have as their *primary* purpose pragmatic needs, and only *secondarily* generation of new knowledge. Here is the heart of the validity versus utility debate. However, there is another way of thinking about the matter. Almost a century ago, A. Lewis[2] noted: "Classifications may be useful for the wrong ends... The clinician may never come to see how vicious are the uses to which he has been, contentedly, putting his classification". Reversing the DSM/ICD view, Lewis held that nosology had to be "valid and useful". If invalid, a nosology is not useful. He concluded: "A valid classification is one which is not only useful, but useful for sound medical and scientific ends". Put another way, the *primary* source for diagnostic classification should be our best scientific knowledge, i.e., classification should be valid scientifically, first and foremost, and also clinically relevant. Only secondarily, in rare cases, can purely utilitarian diagnosis be justified when there is compelling clinical need but zero scientific evidence. DSM/ICD reverses the terms, with hundreds of scientifically unjustified utilitarian diagnoses, versus only a dozen or two with some scientific bases.

This is the kernel of the problem: should validity be central to the diagnostic process, or can we just give up on it, and happily celebrate utility?

To answer this question, let's go back for a history lesson. The original justification for the radical changes of DSM-III in 1980 was that it represented a common language, providing "reliability" and utility. This was not the final goal, though. The claim was made repeatedly that this reliability/utility would be a way-station to validity[3]. In other words, we would get to validity more effectively by having a reliable common language. We would change this language with further scientific research, each revision of DSM moving gradually closer to validity. However, as Jablensky admits, the DSM project has failed to achieve validity. And now we are told that we should change our goal to pure utility, an attempt to make a virtue out of defeat.

Recent debates around DSM-5 have exposed some ideas which previously were expressed mainly behind closed doors. We learned that our DSM leaders have important post-modernist assumptions: they have given up entirely on the whole concept of validity[4].

Contrary to initial DSM-III claims about achieving gradual validity in the future, we now have 40 years of the converse experience. The DSM-IV and 5 leadership stated very explicitly to their task forces that they should make as few changes as possible[4]. This is an anti-scientific attitude. Scientists do not make and test hypotheses by saying to themselves: "Now, let's make as few changes to prior beliefs as possible". DSM classification is now a pure paean to utility, entirely "pragmatic", in the worst meaning of the term: an extremist utilitarianism that has no purpose other than to reflect the wishes and beliefs of the American Psychiatric Association or

DSM leadership, or the loudest interest groups. This statement is documented by historians who have reviewed internal DSM documents[5,6].

Besides its basic anti-scientific attitude, DSM revisions have used higher and higher thresholds for making changes based on research, making it harder to move toward empirically-based validity. Call it the "Sisyphus problem": researchers obtain data, rolling the boulder of knowledge up the hill of ignorance; then DSM leaders say it is not good enough. Another generation of researchers adds to that knowledge, and, if their results pass the DSM task force itself, they are vetoed in the American Psychiatric Association by the Scientific Review Committee, or the Board of Trustees.

We have an unimpeachable example of this Sisyphus problem in the work of the great psychiatric researcher J. Angst. For a century, ever since Kraepelin, the standard view in world psychiatry was that it did not matter if patients were manic (bipolar) or depressed (unipolar), but rather that all mood episodes reflected the same single manic-depressive illness. Angst's Zürich cohort, collected in the early 1960s, suggested that bipolar and unipolar groups differentiated on diagnostic validators of course and genetics[7]. Hence the underappreciated radical anti-Kraepelinian change in DSM-III: the creation of bipolar disorder and major depressive disorder out of Kraepelin's concept of manic-depressive illness. In the intervening decades, with over 40 years of more data, Angst now finds that his Zürich cohort does *not* differentiate well into bipolar and unipolar based on course and other diagnostic validators[8]. The same Zürich dataset, now even more valid with complete prospective follow-up of the entire lifetime of its subjects, is rejected by the DSM-5 task force. What was considered acceptable to make very radical changes in the 1970s is now rejected decades later for even minor changes (like duration of hypomania or definition of mixed states). Much more radical changes in the past were made with much less science.

There is not even a utilitarian justification for this resistance. DSM-5 field trials now indicate that, after four decades, major depressive disorder has poor reliability[9], even worse than in the past. Our current nosology of major depression is *both* false and useless.

Angst, being a true scientist, falsifies his own hypotheses, something the DSM/ICD leadership has been unwilling to do, which brings us to the most baneful consequence of the rejection of science/validity in favor of pragmatism/utility: because of DSM/ICD, all research, both clinical and biological, is doomed to failure. This self-fulfilling prophecy is then used by DSM advocates of pragmatism/utility to justify further their rejection of science-based classification. We reach a dead end in obtaining further new knowledge precisely because obtaining new knowledge is "secondary" to the pragmatism that ensures that no new knowledge will be achieved. Psychiatric progress never occurs, because it cannot occur with these anti-scientific attitudes.

To state it otherwise: DSM/ICD is a "social construction". That's what the concept of utility means. It is created for social – professional, insurance, forensic, economic, ideological, political, cultural – purposes. It is not, as admitted by Jablensky, primarily based on scientific research. The fact that DSM/ICD is a social construction reflects its underlying philosophy, post-modernism[10].

If we create diagnostic categories based on social, economic and political considerations, why should genes correlate with those categories? Why should neuroanatomy correlate with wishes for insurance reimbursement? When DSM/ICD phenotypes for biological studies are purely social constructions, it should be no surprise that hardly any major genes/biomarkers for DSM/ICD diagnoses are identified. Four decades of failure in DSM-based research are hard to ignore. Recent change in the U.S. National Institute of Mental Health (NIMH) policy, such that DSM criteria are no longer acceptable for research[11], is an institutional verification that an emphasis on utility actually *prevents* ever achieving validity.

Because DSM failed, one should not conclude, as the NIMH leadership does, that the whole clinical research project failed. In fact, because of DSM pragmatism, clinical research has *not* been the main basis of our diagnostic system for 40 years. Let us now not draw the false conclusion that clinical research into psychiatric diagnosis has failed, when instead it has been ignored.

Nor will it do to resort to prayer – wishing for a gene, or a brain circuit, that will someday, somehow, split the Red Sea. The gene/biological marker miracle will never happen as long as DSM/ICD fails to put science first[12].

The explicitly vague term "disorder" reflects post-modernist cynicism about the disease concept[13]. The attempt to base "disorder" definitions on functional impairment and severity of symptoms is not conceptually, biologically, or scientifically sound. There are many medical diseases that do not cause functional impairment (such as silent cancers), or involve mild rather than severe symptoms. Some medical diseases even are associated with some benefits, rather than only harms (e.g., decreased malaria risk with sickle cell trait). The extremist DSM/ICD ideology of rejecting mild symptoms does not solve the "false positives problem" nor improve predictive values of diagnosis[14]. Instead, it feeds into, and perhaps reflects, stigma against mental illness, an ironic result of DSM/ICD "pragmatism", understandable as another baneful effect of cultural post-modernism.

In sum, my main critique is that a primarily utilitarian approach, in the end, is not useful, because it matters – cultural post-modernist assumptions notwithstanding – whether we are really right or wrong, i.e., whether our diagnoses are valid. In clinical medicine, where lives are in the balance and where scientific values are accepted, any other view is difficult to defend.

S. Nassir Ghaemi
Mood Disorders Program, Tufts Medical Center, Boston, MA, USA

1. Jablensky A. World Psychiatry 2016;15:26-31.
2. Lewis A. Br Med J 1938;2:875-8.
3. Klerman GL, Vaillant GE, Spitzer RL et al. Am J Psychiatry 1984;141:539-53.
4. Frances A. Pragmatism in psychiatric diagnosis. www.psychologytoday.com.

5. Decker H. The making of DSM-III. New York: Oxford University Press, 2013.
6. Shorter E. Before Prozac: the troubled history of mood disorders in psychiatry. New York: Oxford University Press, 2009.
7. Angst J. Monogr Gesamtgeb Neurol Psychiatr 1966;112:1-118.
8. Angst J. Br J Psychiatry 2007;190:189-91.
9. Regier DA, Narrow WE, Clarke DE et al. Am J Psychiatry 2012;170:59-70.
10. Ghaemi N. On depression: diagnosis, drugs and despair in the modern world. Baltimore: Johns Hopkins Press, 2013.
11. Insel TR. Transforming diagnosis. www.nimh.nih.gov.
12. Ghaemi S. Acta Psychiatr Scand 2014;129:410-2.
13. Ghaemi S. World Psychiatry 2013;12:210-2.
14. Phelps J, Ghaemi SN. Acta Psychiatr Scand 2012;126:395-401.

DOI:10.1002/wps.20287

# We need science to be useful too

Jablensky's notion of a fundamental contrast between utility and validity in psychiatric classification[1] probably bears a relation to the tensions between pragmatic and correspondence ideas of truth. Having both in play at once creates conundrums. Of particular relevance, if one supposes that truth is correspondence with reality as it really is, then mere pragmatic value – utility – will always look like it falls short.

In several key places, Jablensky refers to the shifting nature of utility, contrasted with "reality", understood in this context as inner biological and psychological structure, or "essential structure". There are plenty of places and times in the history of science when it has been reasonably supposed that theory grasped the essential nature of reality. To name but a few: Newtonian mechanics, the mature period table of elements, and the biomedicine of cholera. Also, by way of contrast, in plenty of occasions it did not seem so, such as in the relation between general relativity and quantum mechanics, the models of global warming, the developmental pathology of most medical conditions, the biomedicine of some cancers, and most or all psychiatric conditions.

From a pragmatic point of view, the difference here is a matter of how much the science predicts: in the former kind of case, the theory predicts everything of interest (at the time), while in the latter the theory doesn't at all, or we have a range of sub-theories predicting more or less within sub-domains of interest, but no unified theory. When the idea of truth as correspondence is working in the background, however, the theories which predict everything of interest (at the time) appear as its exemplars, illustrating that our concepts can and therefore should grasp the nature of reality as it really is. In practice, in the sciences, it has become obvious, since the demotion of Newtonian mechanics and the absence of a unified physics, that scientific theories don't stay the same but evolve for many reasons, so it would be rash – misconceived – to say that science grasps reality as it really is, once and for all. We can say that it provides better and better approximations, but this comes down to: it gets better at predicting. Prediction is useful in its own right, but of special interest are predictions that help us solve problems, those that underpin interventions. Science is closely tied to utility and technology.

Psychiatric classification is supposed to have clinical utility. A particular diagnosis is supposed to provide some information useful for clinical management, such as course and prognosis with and without particular treatment(s). By all means diagnoses are only partly successful in this, more or less so depending on the condition, subtype and which treatment. Nevertheless, in the clinic, we suppose that the current diagnostic system guides management somewhat, even if imperfectly, better than nothing, and better than any other system on offer.

Onto this shifting problem domain of clinical utility, Jablensky proposes two criteria of "validity". One of them is that to be valid a condition must be *discrete*, separated from others by a "zone of rarity". This sounds to me like the correspondence theory of truth at work again, because this theory supposes that *facts* and therefore their representations are discrete, each identical to itself and to no other thing. So far as utility is concerned, however, fuzzy overlapping categories can still be useful, more or less, and might be all we have to go on. The weather can be forecast, more or less well, for a limited time ahead, by cloud-shape types (by all means not by shapes of individual clouds), even though not precisely defined and sometimes muddled together.

The other criterion of validity Jablensky proposes is mapping on to the science. He cites the diverse criteria for establishing validity of diagnoses proposed by Robins and Guze, Kendler, and Andreasen. These include, to name but a few, familial aggregation, typical precipitants, psychological tests, neurochemical assays, as well as rates of relapse and recovery, and response to treatment. In these lists, clinical utility appears as validation marker, which, in the view being proposed here, it should, there being no fundamental conceptual distinction between utility and validity. Both utility and validity come to the issue of how much of interest is predicted, and among that, the critical issue of how the predictions guide action and underpin technological solutions.

So what do we expect of scientific validity criteria such as genetic, neurochemical, neurological or neuropsychological? We expect these to be useful too and value them for this reason. We do not expect them just to "map onto reality", otherwise understood. As mentioned earlier, the biomedical model of cholera can be reasonably described as pinning down the real nature of the disease, but this description is underpinned by the fact that the model delivers everything of interest, specifically models of and technologies for treatment and primary prevention.

Increasingly we know that the causes of psychiatric conditions – along with the causes of many general medical conditions – are not singular but multi-factorial, and moreover may have a development from